Benchmark Beispiele (Szenarien mit Zufall)

Glückspiele:

1. Wir werfen ein Würfel. Was ist die Augenzahl?

2. Wir spielen Lotto. Wie viele richtige Zahlen haben wir?

3. Wir spielen ROT im Roulette 100 mals. Wie viel mal gewinnen wir?

4. Wir spielen ROT im Roulette bis wir zum ersten Mal gewinnen. Wie lang müssen wir warten?

Auch ein paar Beispiele, die NICHT Glückspiele sind:

5. Wir telefonieren unseren besten Freund an. Es ist besetzt. Wie lang müssen wir warten, bis er auflegt?

6. Wir besitzen eine Versicherungsfirma. Wie viele Schadensmeldungen erhalten wir zwischen 10^{00}–11^{00} an 5. Januar 2017?

---

1. – 6. sind verschiedene Zufallsexperimente

Wir möchten sinnvolle Modelle definieren, die auch hilfreich sind, zu etwas über die Zukunft lernen.

Diskrete W-räume: Beispiele

(1) **Laplaceraum** (Ω, ℓ), $|\Omega| < \infty$.

(Gleichverteilung) $\ell : \Omega \to [0,1]$

$\forall \omega \in \Omega$ $\quad \boxed{\ell(\omega) = \frac{1}{|\Omega|}}$

• Check: $\sum_{\omega \in \Omega} \ell(\omega) = \sum_{\omega \in \Omega} \frac{1}{|\Omega|} = |\Omega| \cdot \frac{1}{|\Omega|} = 1$ ✓

Beispiele: • faire Münze: $\Omega = \{\text{Kopf}, \text{Zahl}\}$

• fairer Würfel: $\Omega = \{1, 2, 3, 4, 5, 6\}$

(2) **Bernoulliraum** zum Parameter $p \in [0,1]$

(Ω, be_p) $\quad \Omega = \{0, 1\}$ $\quad$ 1: "Erfolg"
0: "Misserfolg"

$be_p(0) = 1 - p$

$be_p(1) = p$

• Check: $\sum_{\omega \in \Omega} be_p(\omega) = (1-p) + p = 1$ ✓

Beispiele: • Wir spielen im Roulette einmal ROT.

$$p = \frac{18}{37}$$

• (nicht unbedingt fairer) Münzwurf

(3) Hypergeometrische Verteilung $\boxed{\text{zu den Parametern } N, n, R \in \mathbb{N}_0,\ 0 \le R, n \le N}$

$$\left(\Omega, h_{N,n,R}\right)$$

$$\Omega = \{0, 1, \dots, n\}$$

$$h_{N,n,R}(k) = \frac{\binom{R}{k}\binom{N-R}{n-k}}{\binom{N}{n}} \qquad \forall k \in \Omega$$

• Check: $\sum_{k=0}^{n} h_{N,n,R}(k) = 1$

HAUSAUFGABE

Behauptung: Urne mit N gefärbten Kugeln (~~[illegible]~~)
$N = R$ rot $+ (N-R)$ blau

n Ziehen, OHNE Zurücklegen

W-keit, genau k rote Kugeln zu ziehen $= h_{N,n,R}(k)$

Beweis: Wir arbeiten in dem Laplaceraum auf $\binom{[N]}{n}$.

Numerieren wir die Kugeln ~~[illegible]~~, ~~[illegible]~~ sagen wir die Kugeln $1, \dots, R$ sind rot und Kugeln $R+1, \dots, N$ sind blau

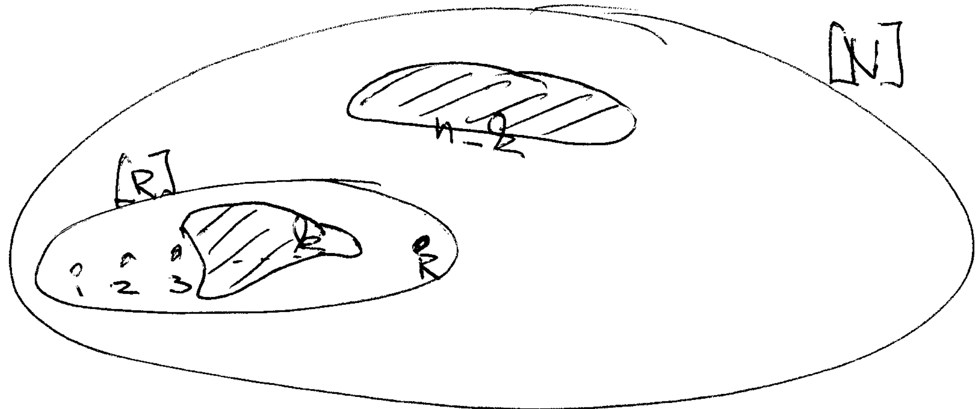
"Günstige" Elementarereignisse: $G_k := \left\{ A \in \binom{[N]}{n} : |A \cap [R]| = k \right\}$

• $[R]$ ist die Menge der roten Kugeln

• $|A \cap [R]|$ ist die Anzahl der roten Kugeln in A

"Alle" Elementarereignisse: $\binom{[N]}{n}$

Was ist $|\mathcal{G}_k|$?



Mehrstufige Produktregel um die Elemente A von $\mathcal{G}_k$ zu zählen.

Frage 1.: Was ist der Schnitt von A und $[R]$?

Da A genau k rote Kugel hat, ist die Anzahl der Möglichkeiten für $A \cap [R]$: $\left|\binom{[R]}{k}\right| = \underline{\underline{\binom{R}{k}}}$

Frage 2: Was ist $A \cap ([N] \setminus [R])$?

Da A genau $n-k$ blaue Kugeln hat, ist die Anzahl der mögliche $(A \cap [N] \setminus [R])$: $\left|\binom{[N]\setminus[R]}{n-k}\right| = \binom{N-R}{n-r}$

(unabhängig von der ersten Antwort.)

- $A \cap [R]$ und $A \cap ([N] \setminus [R])$ bestimmen A eindeutig

$$\rightsquigarrow \quad |\mathcal{G}_k| = \binom{R}{k} \cdot \binom{N-R}{n-k}$$

W-keit in Frage: $\dfrac{|\mathcal{G}_k|}{\binom{N}{n}} = \dfrac{\binom{R}{k}\binom{N-R}{n-k}}{\binom{N}{n}}$ ✓ $\blacksquare$

~~Alternative Beispiel~~ Bem: Abbildung $A \longmapsto (A \cap [N], A \cap ([N] \setminus [R]))$ ist eine Bijektion von $\mathcal{G}_k$ nach $\binom{[R]}{k} \times \binom{[N]\setminus[R]}{n-k}$.

(Inverse (Abbildung) ist: $(M_1, M_2) \longrightarrow M_1 \cup M_2$)

Beispiele (für Hypergeometrische Verteilung)

(1) Lotto $N = 49$, $n = 6$, $R = 6$

- Rote Kugeln sind meine angekreuzten Zahlen
- $n=6$ Ziehen von Lottofirma

Dies war unser Zufallsexperiment No 2. am Anfang.

W-keit von 6-er $= h_{49,6,6}(6) = \dfrac{1}{\binom{49}{6}} = \dfrac{1}{13.983.816}$

5-er $= h_{49,6,6}(5) = \dfrac{\binom{6}{5}\cdot\binom{43}{1}}{\binom{49}{6}} = \dfrac{1}{55{,}492}$

4-er $= h_{49,6,6}(4) = \dfrac{\binom{6}{4}\cdot\binom{43}{2}}{\binom{49}{6}} = \dfrac{1}{1033}$

3-er $= h_{49,6,6}(3) = \dfrac{\binom{6}{3}\cdot\binom{43}{3}}{\binom{49}{6}} = \dfrac{1}{56{,}7}$

2-er $= h_{49,6,6}(2) = \dfrac{\binom{6}{2}\cdot\binom{43}{4}}{\binom{49}{6}} = \dfrac{1}{8{,}3} \approx 0{,}132$

1-er $= h_{49,6,6}(1) = \dfrac{\binom{6}{1}\cdot\binom{43}{5}}{\binom{49}{6}} \approx 0.413$

0-er $= h_{49,6,6}(0) = \dfrac{\binom{43}{6}}{\binom{49}{6}} \approx 0.436$

---

5 von 90 Lotto $\rightsquigarrow h_{90,5,5}$

"Psychologie": Fast jede zweite Woche passiert etwas

5-er $= \dfrac{1}{\binom{90}{5}} \approx 2.28 \cdot 10^{-8}$

4-er $\approx 9{,}7 \cdot 10^{-6}$

3-er $\approx 8{,}1 \cdot 10^{-4}$

2-er $\approx 0{,}022$

1-er $\approx 0{,}23$

0-er $\approx 0{,}74$

5 von 90: nur jede vierte Woche passiert etwas

Beispiel 2. für hypergeometrische Verteilung:

Meinungsbild: Unterstützen Sie Politiker XY? JA / NEIN

$N = 40.000.000$

Rote Kugel: Leute die Unterstützen Politiker XY

$R = ?$ (wir wissen diese Zahl nicht, aber wir möchten eine Idee darüber haben)

$n = 1000$ (wir fragen n zufällige Leute)

Was können wir schließen, wenn wir 750 JA Antworten bekommen haben?
(Und unsere Auswahl der Leute WIRKLICH zufällig war...)

Natürlich: wir möchten so etwas sagen:
"Etwa 30.000.000 Menschen unterstützen Politiker XY"

Aber es ~~kann~~ ist nicht sicher: es kann sein dass in Realität R ist 17.000.000 und wir einfach 750 von unseren 1000 zufällige Kugeln von diesen 17.000.000 gezogen haben.
Dies scheint nicht sehr Wahrscheinlich, denn
$h_{40000000,\,1000,\,17000000}(750)$ ist sehr-sehr-sehr klein.

Aber um in der Lage zu sein, eine Aussage der folgenden Art zu machen: "Die Anzahl der Unterstützer von Politiker XY ist zwischen 29 und 31 millionen mit Sicherheit 98%"
müssen wir die hypergeometrische Verteilung studieren.

- Ist es wirklich wichtig dass die "zufälligen" 1000 Leute, die ich für die Meinungsbild auswähle, alle <u>verschiedene</u> sind? Oder kann ich auch die Leute "mit Zurücklegen" auswählen?

- Wir werden sehen dass, die zwei Meinungsbilder (<u>mit</u> oder <u>ohne</u> Zurücklegen die zufälligen Leute auswählen) zu ähnlichen Ergebnissen führen (weil $n = 1000$ so viel kleiner im Vergleich zu N, R, und $N-R$ ist.)

$\boxed{\text{Binomial Verteilung}}$ zu den Parametern n, p, $n \in \mathbb{N}_0$, $p \in [0,1]$

$$(\Omega, b_{n,p})$$

$$\Omega = \{0, 1, \dots, n\}$$

$$b_{n,p}(k) := \binom{n}{k} p^k (1-p)^{n-k}$$

- <u>Check</u>: $\sum_{\omega \in \Omega} b_{n,p}(\omega) =$

$$= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

$$\underset{\uparrow \text{ Binomialsatz}}{=} (p + (1-p))^n = 1^n = \underline{1}$$

<u>Behauptung</u>: Urne mit N gefärbte Kugel $\quad N, n, R \in \mathbb{N}_0$, $0 \le n, R \le N$

$N = R$ rot $+ (N-R)$ blau

n Ziehen, <u>MIT</u> Zurücklegen

W-keit von <u>genau</u> k rote Kugel zu ziehen $= b_{n, \frac{R}{N}}(k)$

Beweis: Wir numerieren die Kugeln, es seien Kugeln $1, 2, \dots, R$ rot und Kugeln $R+1, \dots, N$ blau.

Laplaceraum auf $[N]^n$

"Günstige" Elementarereignisse: n-Tupel mit genau k Koordinaten die $\le R$ sind (d.h. Kugel ist rot)

Formal $G_k = \left\{ (x_1, \dots, x_n) \in [N]^n : \left| \{ i \in [n] : 1 \le x_i \le R \} \right| = k \right\}$

Wie groß ist $|G_k|$?

Mehrstufige Produktregel:

Frage #1: Welche k der n Koordinaten halten rote Kugeln?

$\binom{n}{k}$ Möglichkeiten

Frage #2: Was ist das k-Tupel von roten Kugeln von links ~~nach~~ nach rechts?

R^k Möglichkeiten (weil jedes der möglichen k-Tupel von $[R]^k$ vorkommen kann)

Frage #3: Was ist das $(n-k)$-Tupel von blauen Kugeln von links nach rechts?

Alle Tupel von $([N] \setminus [R])^{n-k}$ sind möglich

$$\rightsquigarrow \quad |G_k| = \binom{n}{k} \cdot R^k \cdot (N-R)^{n-k}$$

$$\text{W-keit von genau } k \text{ roten} = \frac{|G_k|}{|[N]^n|} = \frac{\binom{n}{k} \cdot R^k \cdot (N-R)^{n-k}}{N^n} =$$

$$= \binom{n}{k} \frac{R^k}{N^k} \cdot \frac{(N-R)^{n-k}}{N^{n-k}} = \binom{n}{k} \left(\frac{R}{N}\right)^k \left(1 - \frac{R}{N}\right)^{n-k} = f_{n,\frac{R}{N}}(k)$$

$\square$

Approximation der hypergeometrische Verteilung durch die binomiale Verteilung

~~Wenn N und R sind groß im Vergleich zu n,~~

~~dann~~ Die genaue Formel für die hypergeometrische Verteilung ist kompliziert:

$$h_{N,R,n}(k) = \frac{\binom{R}{k}\cdot\binom{N-R}{n-k}}{\binom{N}{n}} = \frac{R!\,(N-R)!\;n!\,(N-n)!}{k!\,(R-k)!\,(n-k)!\,(N-R-n+k)!\,N!}$$

$\boxed{\text{Warum?}}$ Denn wir ziehen OHNE zurücklegen.

Für die erste Kugel die Chance für Rot : $\frac{R}{N}$

Aber für die zweite Kugel, die Chance für Rot hängt von dem Ergebnis der ersten Ziehung ab:

Entweder $\frac{R}{N-1}$ oder $\frac{R-1}{N-1}$

Und so weiter: die W-keit für rote in der $i$ten Ziehung hängt von der Anzahl der Erfolge in den vorherigen Runden ab.

$i \leq n$

Es ist immer zwischen $\frac{R-i+1}{N-i+1}$ und $\frac{R}{N-i+1}$

$\boxed{\text{Idee:}}$ Wenn N und R sind groß im Vergleich zu n, dann ~~...~~ $\frac{R-i+1}{N-i+1} \approx \frac{R}{N} \approx \frac{R}{N-i+1}$

$\rightsquigarrow$ W-keit von rot in $i$ten Ziehung $\approx$ W-keit von rot in $i$ste Ziehung MIT ZURÜCKLEGEN

Hoffentlich kann man die W-keit mit der einfacheren "Binomial-

Formal:

<u>Satz</u>: Es seien $k, n \in \mathbb{N}_0$ mit $0 \leq k \leq n$

und sei $p \in (0,1)$.

Für jede Folgen $R_1, R_2, \ldots \in \mathbb{N}$

$N_1, N_2, \ldots \in \mathbb{N}$ von natürlichen Zahlen

so dass $\lim_{l \to \infty} N_l = +\infty$

und

$\lim_{l \to \infty} \frac{R_l}{N_l} = p$

gilt

$$\boxed{\lim_{l \to \infty} h_{N_l, R_l, n}(k) = b_{n,p}(k)}$$

<u>Bemerkung</u>: $\lim_{l \to \infty} R_l = +\infty$ weil $p \neq 0$

$\lim_{l \to \infty} N_l - R_l = +\infty$ weil $p \neq 1$

<u>Beispiel</u>: Für das Meinungsbild, mit $N = 40{,}000{,}000$

$R = 17{,}000{,}000$

$n = 1000$

hoffentlich $\underset{\binom{1000}{k}\left(\frac{17}{40}\right)^k\left(\frac{23}{40}\right)^{1000-k}}{b_{1000,\frac{17}{40}}(k)} \overset{0{,}0071159}{=}$ wäre eine gute Approximation für die genauer W-keit $\overset{0.0071156}{=} \frac{\binom{17000000}{k}\binom{23000000}{1000-k}}{\binom{40000000}{1000}}$

5. Poisson Verteilung

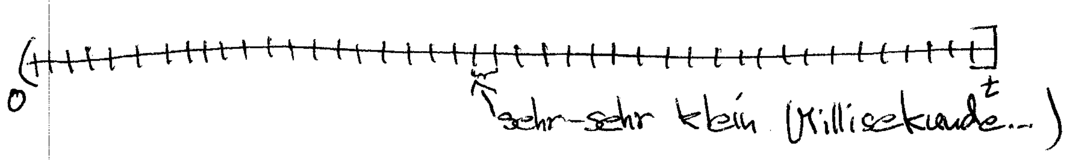
Wie viele Schadensmeldungen erhält z.B. eine Kfz-Haftpflichtversicherung in einem festen Zeitintervall $(0,t] > 0$?

Welcher W-Raum ist vernünftig, dies zu modellieren?

$\Omega := \mathbb{N}_0 \qquad \mathbb{P} = ?$

Heuristische Überlegung:

- Zerlegen wir das Zeitintervall $(0,t]$ in n Teilintervalle der Länge $\frac{t}{n}$



- Annahme: (1) In jedem Teilintervall höchstens ein Schaden eintritt

  (2) W-keit für solch einen Schadenfall ist proportional zur Länge des Teilintervalls $\rightsquigarrow$ $\boxed{\exists \text{ Konstante } \alpha}$, s.d. W-keit eines Schadenfall in einem festen Intervall der Länge $\frac{t}{n}$ ist $\boxed{\alpha \frac{t}{n}}$

  (3) Auftreten eines Schadens in einem Teilintervall nicht davon abhängt, ob in einem anderen ein Schaden auftritt (oder nicht)

$\rightsquigarrow$ "Unglücksfee" zieht für jede der n Teilintervalle aus einer Urne, die einen Anteil $\alpha\frac{t}{n}$ von "Schadenskugeln" enthält.

(Mit Zurücklegen, Ergebnis: geordnet)

$\rightsquigarrow$ W-keit für genau k Schaden $= b_{n,\frac{\alpha t}{n}}(k)$

$\rightsquigarrow$ Für unsere $\mathbb{P}$ wählen wir n größer und größer

$$\mathbb{P}(k) := \lim_{n\to\infty} b_{n,\frac{\alpha t}{n}}(k)$$

<u>Definition:</u> $\boxed{\text{Poisson Verteilung}}$ zum Parameter $\lambda > 0$

(Ω, p_λ)

$\Omega = \mathbb{N}_0$

$\forall k \in \Omega \quad p_\lambda(k) := \frac{\lambda^k}{k!} e^{-\lambda}$

- <u>Check:</u> $\sum_{\omega \in \Omega} p_\lambda(\omega) = \sum_{k=0}^{\infty} p_\lambda(k)$

$$= \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \underbrace{\left(\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \right)}_{= e^{\lambda}} = 1$$

Approximation der Binomialverteilung durch die Poissonverteilung

<u>Satz:</u> Es sei $\lambda > 0$. Dann für ~~alle~~ $\forall k \in \mathbb{N}_0$

$$\lim_{n \to \infty} b_{n, \frac{\lambda}{n}}(k) = p_\lambda(k)$$

<u>Bemerkung:</u> Gleiche Folgerung gilt auch, wenn, statt $p_n = \frac{\lambda}{n}$, p_n : eine beliebige Folge in $[0,1]$ mit $\lim_{n \to \infty} n p_n = \lambda$ ist.

Dann: $\lim_{n \to \infty} b_{n, np_n}(k) = p_\lambda(k)$.

Beweis: Immer erinnern: k ist <u>Konstant</u>, ändert sich NICHT wenn n nach ∞ läuft.

$p_n = \frac{\lambda}{n}$

$$b_{n,\frac{\lambda}{n}}(k) = \binom{n}{k} \cdot p_n^k (1-p_n)^{n-k} = \frac{n(n-1)\cdot\ldots\cdot(n-k+1)}{k!} \cdot p_n^k \cdot \frac{(1-p_n)^n}{(1-p_n)^k}$$

$$= \frac{1}{k!} \frac{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)}{n\cdot n\cdot\ldots\cdot n} \cdot n^k \cdot p_n^k \cdot \frac{(1-p_n)^n}{(1-p_n)^k}, \qquad (1-p_n)^n \to e^{-\lambda}$$

$\frac{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)}{n\cdot n\cdot\ldots\cdot n} \xrightarrow{n\to\infty} 1^k$ weil k constant ist und $\frac{n-i}{n} = 1 - \frac{i}{n} \to 1$ für $0 \le i \le k-1$

$n^k \cdot p_n^k = (n\cdot p_n)^k \xrightarrow{n\to\infty} \lambda^k$

$(1-p_n)^k \xrightarrow{n\to\infty} (1-0)^k = 1$ weil k konstant ist

~~[illegible]~~ $(1-p_n)^n = \left(1-\frac{\lambda}{n}\right)^n \xrightarrow{n\to\infty} e^{-\lambda}$ weil λ Konstant ist.

<u>Also:</u>

$$b_{n,\frac{\lambda}{n}}(k) \xrightarrow[n\to\infty]{} \frac{1}{k!} \cdot 1^k \cdot \lambda^k \cdot \frac{e^{-\lambda}}{1^k} = \frac{\lambda^k}{k!} e^{-\lambda} = p_\lambda(k)$$

$\square$

<u>Bemerkung:</u> wenn wir nur annehmen, dass $\lim_{n\to\infty} n p_n = \lambda$

Dann benutzt man: $\forall x \in [-0.1, 0.1]$

$$e^{-x-x^2} \le (1-x) \le e^{-x} \Rightarrow e^{-p_n n - p_n^2 n} \le (1-p_n)^n \le e^{-p_n n}$$

$e^{-p_n n - p_n^2 n} \xrightarrow{n\to\infty} e^{-\lambda-0}$, $\quad e^{-p_n n} \xrightarrow{n\to\infty} e^{-\lambda}$

Immer

$\forall$ kleine y: $e^{-y} < 1 - y + \frac{y^2}{2}$

$$\Rightarrow e^{-x-x^2} < 1-(x+x^2) + \frac{(x+x^2)^2}{2} = 1 - x - \frac{x^2}{2} + x^3 + \frac{x^4}{2} < 1-x$$

Zurück zu Benchmark-Beispielen (Zufallsexperimente)

(1.) Augenzahl bei einem Würfewurf ⇝ Laplaceraum

(2.) Anzahl der richtige Zahlen beim Lotto ⇝ Hypergeometrische Verteilung

(3.) Anzahl der ROT während 100 Roulette-runde ⇝ Binomialverteilung

(6.) Anzahl der Anrufe während einem festen Zeitintervall ⇝ Poissonverteilung

---

(4.) Wie lang müssen wir warten für das erste ROT im Roulette?

$\boxed{\text{Geometrische Verteilung}}$ zum Parameter q, $0<q<1$

$\Omega = \mathbb{N}$

$\forall k \in \Omega \quad g_q(k) := q^{k-1}(1-q)$

Check: $\displaystyle\sum_{k\in\Omega} g_q(k) = \sum_{k=0}^{\infty} q^{k-1}(1-q) = (1-q)\underbrace{\sum_{j=0}^{\infty} q^j}_{\text{geometrische Reihe}} = (1-q)\cdot\frac{1}{1-q} = 1$

Bemerkung: • q ist die "Misserfolgs"-Wahrscheinlichkeit

• Antwort für Zufallsexperiment #4: Verteilung $g_{\frac{19}{37}}$

• Die Verwendung der Formel $q^{k-1}(1-q)$ für jede k ist voll gerechtfertigt, (die Anzahl der k-Tupel von Kugeln aus einer Urne mit einem Anteil von $(1-q)$ rote Kugeln sodass die ersten $k-1$ blau sind und die letzte ist rot, IST $(qN)^{k-1}\cdot(1-q)N$
ALLE MÖGLICHE $= N^k$. Aber der RAUM $([N]^k)$ ist verschieden für jedes k.